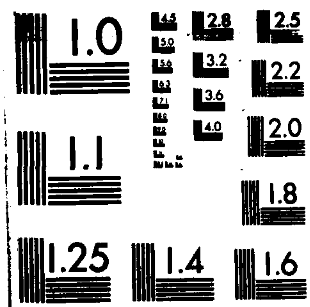
AD-A156 229 MANUAL AND COMPUTER-AIDED SEQUENTIAL DIAGNOSTIC INFERENCE(U) SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION INC ST CLOUD FL S E GORDON JUN 85 AFHRL-TP-84-51

1/1

UNCLASSIFIED F/G 9/2 NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFHRL-TP-84-51

(7)

AIR FORCE

AD-A156 229

HUMAN RESOURCES

MANUAL AND COMPUTER-AIDED SEQUENTIAL DIAGNOSTIC INFERENCE

Sallie E. Gordon

Department of Psychology
University of Idaho
Moscow, Idaho 83843

LOGISTICS AND HUMAN FACTORS DIVISION
Wright-Patterson Air Force Base, Ohio 45433-6503

June 1985
Final Technical Paper for Period May 1984 - August 1984

Approved for public release; distribution unlimited.

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

85 6 10 015

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Public Affairs Office has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

DONALD C. TETMEYER, Colonel, USAF
Chief, Logistics and Human Factors Division

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFHRL-TP-84-51
6a. NAME OF PERFORMING ORGANIZATION Southeastern Center for Electrical Engineering Education 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Logistics and Human Factors Division
6c. ADDRESS (City, State, and ZIP Code) Central Florida Facility 1101 Massachusetts Ave St Cloud, Florida 32769	7b. ADDRESS (City, State, and ZIP Code) Air Force Human Resources Laboratory Wright-Patterson Air Force Base, Ohio 45433-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Office of Scientific Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-82-C-0035
8c. ADDRESS (City, State, and ZIP Code) Bolling Air Force Base Washington, DC 20332	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO.: PROJECT NO.: 9983 TASK NO.: 04 WORK UNIT ACCESSION NO: 51

11 TITLE (Include Security Classification)
Manual and Computer-Aided Sequential Diagnostic Inference

12 PERSONAL AUTHOR(S)
Gordon, Sallie E.

13a TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 25 May 84 TO 20 Aug 84	June 1985	18

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	automation; inference task
05	09		command and control systems; predictive models

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This paper describes a pilot study on how human subjects process information during a diagnostic inference task. The objective was a descriptive/predictive model of the inference task and how that task could be affected by implementation of an automated system. The study directly supported research being conducted by AFHRL on quantitative techniques to predict the impacts that automation may have on operator performance, by defining its interaction with the operator's information processing (Modelling Impacts of Automation on Non-Automated Tactical Command and Control (C^2) Systems). The pilot study involved testing human subjects who had to infer the identity of two fictitious diseases by sampling up to eight symptom dimensions. A set of process and performance variables were selected for measurement. Signal detection theory served as the data collection design. Results were in line with anticipated outcomes (i.e., certainty increased as more cues were sampled); however, certainty rate of increase was highest for trials where subjects sampled four cues and lowest for trials where subjects sampled eight cues (total number of cues was eight). The pilot study helped formulate a list of critical variables expected to affect the operator's information processing and defined plausible relationships between those processes and automation assistance.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Nancy A. Perrigo Chief, STINFO Office	22b. TELEPHONE (Include Area Code) (512) 536-3877 22c. OFFICE SYMBOL AFHRL/TSR

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

Unclassified

AFHRL Technical Paper 84-51 June 1985

MANUAL AND COMPUTER-AIDED SEQUENTIAL DIAGNOSTIC INFERENCE

By

Sallie E. Gordon

Department of Psychology
University of Idaho
Moscow, Idaho 83843

LOGISTICS AND HUMAN FACTORS DIVISION
Wright-Patterson Air Force Base, Ohio 45433-6503

Reviewed and submitted for publication by

Bertram W. Cream
Chief, Ground Operations Branch

SUMMARY

It is becoming increasingly obvious that computerized automation can be a useful aid for a wide variety of positions in the command and control network where many of the tasks involve situational assessment or "diagnostic inference." To optimally combine human talent and computer-aiding systems, one must know how the human operator performs the task unaided (and under what circumstances), what subtasks can be allocated to the machine, and what variables affect operator acceptance of the aiding system. This paper presents a theoretical model of the human performance of a diagnostic inference task when unaided by machine, including the variables affecting those inference processes; and a preliminary model of how a computer-aiding system might be expected to fit into the diagnostic system.

PREFACE

The objective of this research was a theoretical model of how human subjects perform a diagnostic inference task unaided by automation. In addition, a preliminary model of how a computer-aiding device might be expected to fit into the diagnostic system was conceptualized.

This study supported research being conducted in WU 3017-08-06, Modelling Impacts of Automation on Non-Automated Tactical Command and Control (C^2) Systems. This project concerns the prediction of changes in cognitive performance as a function of various kinds of automation. The methodology to be developed will assist planners in designing future automated systems that will optimize human performance. Technical issues confronting this research concern the selection or development of quantitative models that can accurately depict human cognitive process and performance, and the tools and techniques which can capture the higher-level interests between operator and automation.

The results of the study described in this paper included a list of variables anticipated to affect human performance in an inference task and the identification of a candidate technique that can be used to measure the effects that these variables may have on the inference task.

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for making possible a very interesting and rewarding Summer Fellowship at the Air Force Human Resources Laboratory, Wright-Patterson AFB, Ohio. She would like to express sincere appreciation to the Laboratory, in particular the Ground Operations branch, for an exceptional working environment. Finally, she would like to thank Rosemarie Preidis for her enormous support and collaboration in a difficult undertaking, Bertram Cream for his insightful administrative guidance, Larry Reed for many helpful discussions, and the other members of the branch for their professional support and friendship.

TABLE OF CONTENTS

LIST OF FIGURES

MANUAL AND COMPUTER-AIDED SEQUENTIAL DIAGNOSTIC INFERENCE

I. INTRODUCTION

It is becoming increasingly obvious that computerized automation can be a useful aid in a wide variety of positions in the armed services. This is especially true in the world of Command and Control (C^2) where much of the work involves complex situational assessment or "diagnostic inference." As technological complexities increase, human operators will have a more difficult time trying to understand, integrate, and utilize the information made available to them. In contrast to man's limited cognitive capacities and well-documented biases,[1,2] a computer can utilize and aggregate large volumes of information using predetermined optimal strategies that are most appropriate for the situation at hand. It is no longer a question of whether computer aiding will be used, but how it will be used.

Just as there are problems inherent in using a completely "manual" system to perform a function, there are also problems in using a completely "automated" system. These problems have been discussed at length elsewhere;[2,3] but let it suffice to say that at the current time, expert systems are not sufficiently advanced to make automated systems infallible or able to deal with the multitude of unforeseen occurrences that are likely in the C^2 environment.

Since neither man nor machine is solely capable of performing situational assessment functions, the solution lies in using both together and relying on the strengths of each (hopefully also minimizing the weaknesses of each). To integrate man and machine successfully for a given task, one must understand how the human perceives and performs the task, and how the machine can be programmed to perform the task (or parts of the task), and then analyze the best way to fit the two together. In the procurement cycle, a common method for developing a computerized aiding system is to intuitively develop a software system that <u>seems</u> as though it could do the job. Little attention is paid to analysis of the entire task and which subtasks could be best performed by the man and which are best left to the machine (a few exceptions do exist). Consideration is usually not given beforehand to how the operators will react to the aiding system nor to what variables will lead to their acceptance or rejection of the new system. Instead, an automated system is designed and a prototype built. Any modifications necessary to make the system compatible with the operator are usually done after this point. This leads to only those changes that seem absolutely necessary and the result is an overall man-machine system that is much less effective and efficient than what could have been achieved.

Part of the problem outlined above results from an inadequate knowledge concerning three vital questions: (a) How does the human operator perform the assessment task when unaided by automation? (b) What subtasks are best performed by the operator, and what subtasks are best performed by the automation device? (c) What factors determine operator acceptance and use of the automated system? The first question (how the operator performs the task) may seem unnecessary to some. However, this information is needed because it directly affects the answers to the second and third questions. That is, if we know <u>how</u> the operator performs the task (<u>not</u> how his/her performance differs from some theoretically optimal strategy), we can determine specifically what capabilities he/she has that we want to preserve in determining the optimal man-machine subtask allocation. In addition, one can argue that how the operator performs the task will largely affect his/her acceptance of the automation. If the automation is extremely different from or incompatible with the operator's way of perceiving and accomplishing the task, then he/she will be less likely to accept and use that automation.

The theories and methodologies of cognitive psychology can be brought to bear on this problem. By mapping out the cognitive processes or strategies that are used by the perceiver under various situational constraints, we can then measure how those processes change as a function of providing a computer-aiding system.

II. OBJECTIVES

The primary objective of this effort was to develop a predictive model of a diagnostic inference task and how that task would be affected by implementation of an automated system. This objective included the following specific goals:

1. Develop a descriptive model of an inference task that is representative of inference tasks in the C^2 system and would be amenable to laboratory research.

2. Determine appropriate techniques for measuring process and performance in the inference task.

3. Determine a preliminary set of independent variables expected to affect the inference process (and performance).

4. Develop a predictive model of the effects of automation on the operator's inference processes.

Accomplishing these goals would serve two purposes: provide guidance to researchers at the Air Force Human Resources Laboratory concerning variables of critical interest in related field research and provide a framework for follow-on laboratory research designed to answer some of the questions outlined earlier.

III. DESCRIPTION OF THE INFERENCE TASK

Diagnostic inference will be defined as the process of using available cues to determine the underlying or "unseen" cause of those cues. An example is medical diagnosis where the doctor must infer a disease that causes some set of cues (symptoms). If the available cues are very informative, the inference will be accurate and made with a high degree of confidence. However, it is often the case that the cues do not convey enough information, and the inference task takes place amid psychological uncertainty.

In the past, most research addressing this type of task assumed a "single-stage" process, wherein the perceiver received the cues and somehow aggregated or operated upon the information and derived a judgment. This was a popular view for some time, partially because it was amenable to laboratory experimentation and formal mathematical description and analysis.[4] Two approaches were common: The first was to develop a formal mathematical model (such as Bayes Theorem) to specify optimal performance and then to fit that model to data obtained with human subjects;[5] the other approach was to use linear regression models to assess how the subjects were combining or utilizing the cues in generating the inference.[6,7] Research questions addressed in these latter studies included such topics as, what cues are predominantly utilized by people,[7] how many dimensions or cues are used for various tasks and whether these cues are same as in the "real world,"[8] and whether experts cluster or weigh cues in a similar fashion.[9,10]

The appropriateness of these models has been debated recently; therefore, this issue will not be discussed at length. However, two points will be reviewed. The first is the criticism that most laboratory inference tasks involve simultaneous and orthogonally manipulated cues. This cue independence is seen as being highly artificial and unrealistic.[11] Since humans develop cognitive skills to deal with a real and complex world, it is not surprising that they perform "suboptimally" on these inference tasks where no intercue correlations are preserved.

The second criticism with these approaches suggests that the inference task should be treated not as a single-stage process but as a multiple-stage process.[12] This is not to say that the reception and the integration of cues are different stages, but that the acquisition of cues or characteristics takes place over time and that this process should be reflected in the theoretical models.

In line with these criticisms and recent views in cognitive psychology, it will be assumed that the inference task of interest takes place in a complex situation where the perceiver must sequentially seek information to make the inference judgment. In addition, that information is typically incomplete and varies in its diagnosticity. The perceiver starts with one or two cues and then searches for others either to confirm hypothesized causes or to suggest new ones. The inference process is viewed as a "constructive" process, much like building a jigsaw puzzle. One does not need all of the pieces to be able to infer the nature of the picture; instead, the ability to draw the inference will depend on the combined information provided by the pieces put together.

In psychological terms, the perceiver uses both Conceptually-Driven processing (where the hypothesized cause suggests cues to seek) and Data-Driven processing (where cues suggest plausible hypotheses). The cyclic procedure continues until the perceiver exceeds some certainty criterion that he/she knows the identity of the cause. In some cases, a lack of information will prevent that criterion from being reached at all.

IV. MEASURING PROCESS AND PERFORMANCE

It was suggested that the inference judgment is constructed over time as information is acquired. This implies that it is important to measure the process by which the perceiver is coming to a conclusion, as well as to measure performance per se. Each of these issues will be addressed in turn.

A. PROCESS

Several methodologies for measuring judgment or decision "process" have been suggested. Payne[13] is a predominant supporter of two of these methods known as process tracing. The first method is a class of measurement techniques where the subject's information acquisition is monitored. The subject must view or select information in a way that can easily be observed and recorded. Data are obtained concerning what cues the subject samples, in what order, how many are sampled, and the amount of time for the cue sampling.

The other method of process tracing is the collection of verbal protocol. In this technique, the subject is simply asked to "think out loud" while performing the task. Although this type of data can give insight to the subject's strategies, it cannot be assumed that the subject will always verbalize the cognitive processes as they occur.

After assessing the various process measurement techniques, a method was decided upon which seemed most suitable to an inference task. The inference is actually a classification task, where the perceiver must choose between class A, class B, class C, and so on (also possibly "none of the above"). The process measure being suggested consists of two aspects:

1. Allowing the subjects to acquire whichever cues they desire until they feel reasonably confident in their choice (this is similar to the previously described information acquisition measure).

2. Asking the subject after each cue acquisition to give the hypothesized cause(s), along with a subjective certainty rating (i.e., 1 = not certain at all, 7 = extremely certain). An example of these measures will be presented shortly.

B. PERFORMANCE

In addition to studying cognitive processes or strategies engaged in by the perceiver, it is informative to determine how well the operator is able to infer the cause of the cues. The most appealing measures are suggested by Signal Detection Theory (SDT) because it allows for separate measurement of discrimination capabilities and subjective bias.[14] However, standard SDT measures cannot at this time be applied to more than a two-category (Signal-Noise) task. Swets and Pickett[15] discuss the problem of multiple-category discrimination and suggest using Percent Correct as a reasonable solution. This is justified because the inference task for multiple causes is actually conceptually similar to a forced-choice task. It was therefore determined that Percent Correct could be used as a performance measure in the diagnostic inference task.

To summarize the inference task and associated measures of process and performance:

1. The subject is given an initial cue.

2. The subject verbalizes one or more hypotheses, along with a certainty rating.

3. The subject samples a new cue of his/her choice.

4. The subject verbalizes revised hypothesis(es), along with new certainty rating.

•
•
•

(continues until subject exceeds some subjective certainty criterion and chooses to stop)

An example of data collected from a subject is given below:

	Class A	Class B	Neither
Cue 1	1		
Cue 2	2	2	
Cue 3		4	
Cue 4		7	

It can be seen that after the first cue sampling, the subject hypothesized class A as the cause, with a certainty rating of 1. However, after receiving additional cues, the hypothesized cause was switched to class B, with increasing certainty.

To determine performance (Percent Correct) from the data, a cutoff point must be chosen for the certainty scale. For example, correct might be arbitrarily defined as "more than 4." If the correct answer for the example was class B, the subject would be scored as "correct" on only the last observation given above. The advantage of this method is that accuracy of the inference can be assessed without the subject's having to make a strictly Yes-No decision. (This is a way of "setting" the subject's decision criterion.) With this method of data collection, it is possible to measure a variety of process and performance variables as the subject samples the information:

1. Information sampled.

2. Hypotheses generated.

3. Hypothesis transition (at what point they give up and generate another).

4. Subjective certainty criteria.

5. Accuracy.

C. FEASIBILITY STUDY

A pilot study was conducted during the Summer Fellowship period to determine the appropriateness of the above variables. Eight subjects (students at Wright State University, Dayton, Ohio) received course credit for participating in the study.

Subjects were first given fictitious names of two diseases, along with eight case studies for each disease. The case studies were described in terms of patient initials and occupations and a set of symptoms that varied in number from three to seven. There were eight total symptom dimensions (such as blood pressure, weight loss, etc.). Four of the dimensions were strongly associated with disease A, and the other four were strongly associated with disease B. However, there was some overlap of symptoms across diseases. After subjects reported that they were familiar with the disease characteristics (approximately 15 minutes), they were given 32 new case studies to diagnose. For each one, they were presented with a 3x5 index card with initials, occupation, and one symptom. They were also presented with seven other cards with a symptom dimension (e.g., "blood pressure") labelled on one side. They were told to turn any card over that they wished, and after each one, to tell the observer their hypothesis and certainty rating. Data were collected in a manner corresponding to the example shown previously.

Based on the theoretical considerations outlined earlier, two effects were expected:

1. The certainty ratings would start low and slowly increase until some criterion was reached; at this point, subjects would discontinue information acquisition.

2. Percent Correct would be a positive function of the number of symptoms sampled on a given trial.

Data relevant to the first hypothesis are presented in Figure 1. It can be seen that certainty ratings did, in fact, start relatively low and increase over cue samples; the rate of increase was highest for trials where subjects sampled only four cues, and lowest for trials where subjects sampled eight cues. Notice that for trials where all symptoms were utilized, the certainty rating never reached a level equal to those for the other trials. That is, a subjective certainty criterion of approximately 5 was exceeded in all but those trials where subjects simply did not have enough information to generate an inference with a high degree of confidence.

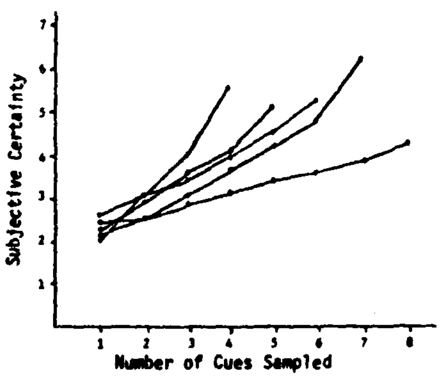


Figure 1. Subjective Certainty as a Function of Cues Sampled.

This pattern in the Certainty data is consistent with the Accuracy (Percent Correct) data shown in Figure 2. A liberal cutoff was arbitrarily chosen so that "3 and above" for the Certainty rating was considered a "correct" answer. Still, the scores were not remarkably high (mean % correct for all trials was .66). This would indicate that subjects were not able to completely learn the correct structure of the disease-symptom associations. Performance varied widely from subject to subject, with a range of .50 to .97 for the eight subjects.

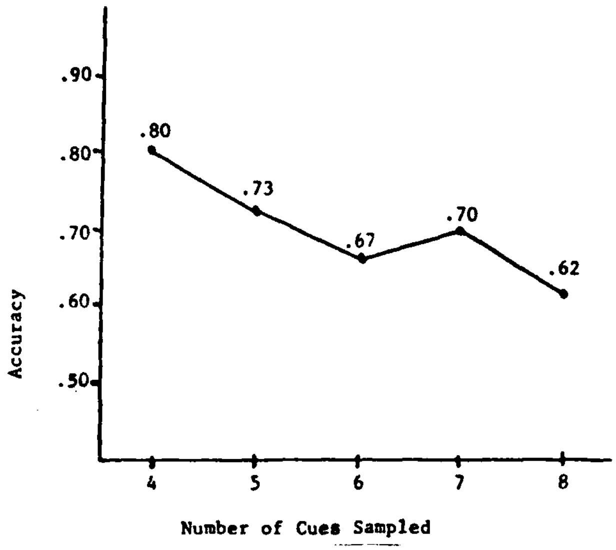


Figure 2. Accuracy as a Function of Cues Sampled.

Figure 2 shows that subjects' perceived uncertainty had some basis, in that performance decreased as the subjects utilized a greater total number of symptoms. Interestingly, both data sets show the same pattern of an increase in accuracy and subjective certainty for trials where seven cues were sampled. Additional studies will be conducted to determine whether this effect holds true for larger sets of data.

It can be seen from the data collected thus far that the process and performance measures are appropriate for the task and yield a rich variety of information concerning the perceiver's strategy. (It should be noted that additional, more fine-grained analyses are planned.)

V. VARIABLES AFFECTING THE INFERENCE PROCESS

Inference process and performance are each affected by various situational constraints. These include characteristics of the perceiver, of the task, and of the situational environment from moment to moment. A real-world illustration of this complex situation will be described as a way of introducing situational constraints which affect process and performance variables.

An important inference function within the C^2 network is that of the radar operator (and/or officer) who must determine the identity of aircraft showing up on the radar scope. This person receives a "track" on the scope and must "infer" the identity of the aircraft, using auxiliary pieces of information or "cues." These pieces of information include Flight Plan Data; Special Codes emitted by the aircraft (friendly aircraft); speed, heading, electronic emissions, and intelligence data; and possibly, visual identification information. Some of this information will be quite diagnostic (e.g., the Codes), while other data will not be particularly diagnostic (speed or heading). Given enough time, the operator could identify almost any track. If nothing else, the operator could send someone up to look at the aircraft. The problem is based on the fact that the operator officially has a maximum of 2 minutes to identify the aircraft. In addition, sending someone up for visual identification is very costly. In wartime conditions, the operator will have to identify many tracks in a very snort period of time.

A preliminary list of variables was developed that are considered most important in determining the operator's performance (as defined by accuracy). These situational constraints on performance are shown in Figure 3; for example, "Time Stress" is a variable that will negatively impact performance. Time Stress will be determined by the number of tracks the operator has to identify, how far away the aircraft are from important locations, and the alert status at the time (i.e., white, yellow, or red alert).

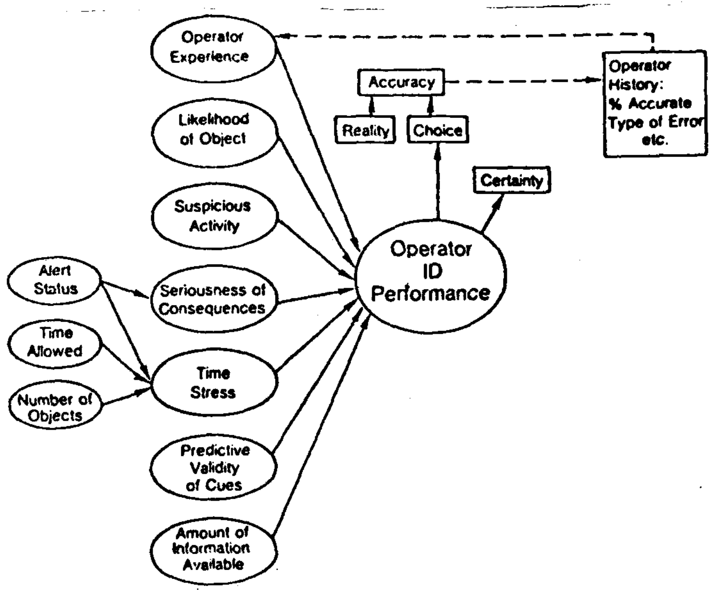


Figure 3. Variables Affecting Operator Performance.

The same variables are assumed to also impact the process of generating an inference. Time stress will result in fewer cues considered and a lower subjective certainty criterion (point where the perceiver is willing to stop collecting data). Time stress is also expected to decrease the number of hypotheses generated and considered during the inference process.

Suspiciousness of the aircraft travel will increase the likelihood of an enemy aircraft being hypothesized, resulting in cues being sampled which will confirm (or disconfirm) that hypothesis. Diagnosticity or predictive validity of the cues will result in both a need to sample fewer cues and a higher subjective certainty concerning the inference. Finally, operator experience is assumed to enhance the inference performance, but it is not totally clear how that variable will impact the process.

VI. IMPACT OF AUTOMATION

Before discussing the impact of implementing a computer-aiding system, it is necessary to define the nature of the automated system. Several researchers have developed taxonomies of types of automation;[16] most are admittedly imprecise. For the present purposes, a continuum will be assumed, with a completely manual method of task accomplishment at one end and complete automation at the other end of the continuum. An example of this type of continuum is shown in Figure 4. One of the more prevalent types of automation is given on the top right of the scale; that is, the computer completely performs the task and provides an "answer" to the operator, who then decides if he/she wants to believe and utilize that answer. This type of automation is currently being planned for the radar identification task described earlier.

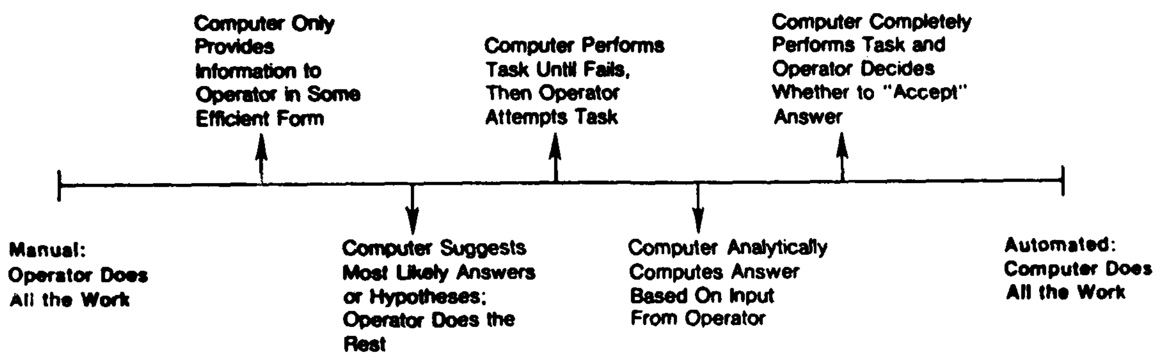


Figure 4. Degrees of Automation.

A predictive model can be outlined at this point, describing the impact of implementing this type of automation. First, we can say that most of the variables expected to affect the human perceiver will probably not affect the performance of the automated system. Thus, under conditions of stress, the human performance will deteriorate, whereas machine performance will not. The only variables expected to affect machine performance will be the characteristics of the cues themselves (see Figure 5). This puts the operator in the place of deciding whether to "trust" the machine, knowing that the machine can perform the task more quickly and objectively in times of duress. The operator's decision to use the answer provided by the system will depend on how much time he/she has, the seriousness of the consequences, and the nature of the information (e.g., if it is a suspicious activity). In addition, the operator's decision to use the answer provided will strongly depend on the operator's feelings about his/her own ability versus the machine's history of reliability and accuracy. If the machine has a relatively low "hit rate," the operator will be more inclined to consider its answer worthless and go mostly on his/her own judgment.

An interesting point which should be considered is the possibility that the operator might treat the answer provided by the machine as merely "another cue." Rather than determining his/her inference based on the standard set of cues and then comparing this with the automated choice, the operator may simply use that choice during the inference process, treating it as he/she does all of the other pieces of information.

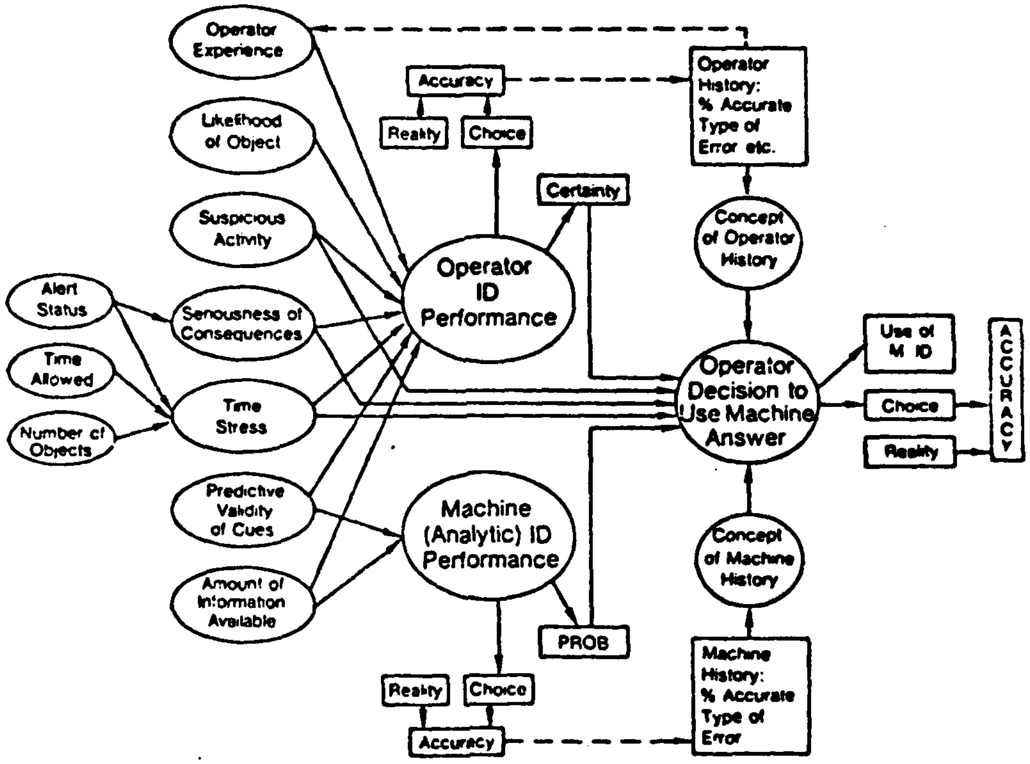


Figure 5. **Implementation of Automation in Inference Task.**

VII. RECOMMENDATIONS

The previous section outlined hypothesized effects of situational constraints on inference process and performance. In addition, a preliminary predictive model was constructed relating the implementation of an automated subsystem within a "manual" inference system. The relationships between the variables need to be obtained empirically to confirm the theoretical assumptions.

I propose to do this by developing a laboratory inference task that has the characteristics outlined earlier (complex, sequential cue acquisition, etc.). This inference task will be run on a computer terminal so that the subject can request cues and respond with hypotheses and certainty ratings. This will allow collection of data relevant to the process and performance variables outlined previously.

For assessment of the processes used in a manually performed inference task, situational constraints will be manipulated and their effects on process and performance variables measured. This will allow determination of what processes are utilized under various task conditions.

A second phase of the follow-on research will be to provide the subject with a computer aid. This will be a simulated computer-aiding system because the answer to be given to the subject will be predetermined by the Experimenter. In this way, the accuracy of the answers provided to the perceiver can be manipulated as an independent variable. It is expected that as accuracy of the automated system increases, the use of the machine will also increase in an exponential fashion. Situational constraints will also be varied, similar to those used in assessment of the completely manual inference process.

REFERENCES

1. Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. Annual Review of Psychology, 28, 1-39.

2. Price, H.E., Maisana, R.E., & Van Cott, H.P. (1982). The allocation of functions in man-machine systems: A perspective and literature review (NUREG CR-2623). Oak Ridge, TN: Oak Ridge National Laboratory.

3. Yntema, D.B., & Torgerson, W.S. (1961). Man-computer cooperation in decisions requiring common sense, IRE Transactions on Human Factors in Electronics, 2, 20-26.

4. Einhorn, H.J., & Hogarth, R.M. (1981). Behavioral decision theory: Processes of judgment and choice, Annual Review of Psychology, 32, 53-88.

5. Slovic, P., & Lichtenstein, S. (1971). Comparison of bayesian and regression approaches to the study of information processing in judgment, Organizational Behavior and Human Performance, 6, 649-744.

6. Einhorn, H.J., Kleinmuntz, D.N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment, Psychological Review, 86, 465-485.

7. Schustack, M.W., & Sternberg, R.J. (1981). Evaluation of evidence in causal inference, Journal of Experimental Psychology: General, 110, 101-120.

8. Ebbeson, E.B., & Konecni, V.J. (1980). On the external validity of decision-making research: What do we know about decisions in the real world? In T.S. Wallston (Ed.), Cognitive processes in choice and decision behavior. Hillsdale, NJ: Lawrence Erlbaum.

9. Einhorn, H.J. (1974). Expert judgment: Some necessary conditions and an example, Journal of Applied Psychology, 3, 562-571.

10. Ashton, R.H. (1974). Cue utilization and expert judgments: A comparison of independent auditors with other judges, Journal of Applied Psychology, 59, 437-444.

11. Medin, D.L., Altom, M.W., Edelson, S.M., & Freko, D. (1982). Correlated symptoms and simulated medical classification, Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 37-50.

12. Elstein, A.S., Shulman, L.E., & Sprafka, S.A. (1978). Medical problem solving: An analysis of clinical reasoning. Cambridge, MA: Harvard University Press.

13. Payne, J.W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis, Organizational Behavior and Human Performance, 16, 366-387.

14. Green, D.M., & Swets, J.A. (1966). Signal detection theory and psychophysics. New York: Wiley.

15. Swets, J.A., & Pickett, R.M., (1982). Evaluation of diagnostic systems: Methods from signal detection theory. Cambridge, MA: Bolt, Beranek and Newman Inc.

16. Hart, S.G., & Sheridan, T.B. (1984). Pilot workload, performance, and aircraft control automation, Proceedings of the AGARD Symposium on Human Factors Considerations in High Performance Aircraft. Williamsburg, VA.

ATE

LMED

-8